

PROCEEDINGS OF SPIE

SPIDigitalLibrary.org/conference-proceedings-of-spie

Cochlear shape description and analyzing via medial models

Johannes Gaa, Lüder A. Kahrs, Samuel Müller, Omid Majdani, Tobias Ortmaier

Johannes Gaa, Lüder A. Kahrs, Samuel Müller, Omid Majdani, Tobias Ortmaier, "Cochlear shape description and analyzing via medial models," Proc. SPIE 9413, Medical Imaging 2015: Image Processing, 941345 (20 March 2015); doi: 10.1117/12.2082033

SPIE.

Event: SPIE Medical Imaging, 2015, Orlando, Florida, United States

Cochlear Shape Description And Analyzing Via Medial Models

Johannes Gaa^a, Lüder A. Kahrs^a, Samuel Müller^a, Omid Majdani^b, Tobias Ortmaier^a

^aLeibniz Universität Hannover, Institute of Mechatronic Systems, Hanover, Germany

^bHannover Medical School, Department of Otolaryngology, Hanover, Germany

ABSTRACT

Planning and analyzing of surgical interventions are often based on computer models derived from computed tomography images of the patient. In the field of cochlear implant insertion the modeling of several structures of the inner ear is needed. One structure is the overall helical shape of the cochlea itself. In this paper we analyze the cochlea by applying statistical shape models with medial representation. The cochlea is considered as tubular structure. A model representing the skeleton of training data and an atomic composition of the structure is built. We reduce the representation to a linear chain of atoms. As result a compact discrete model is possible. It is demonstrated how to place the atoms and build up their correspondence through a population of training data. The outcome of the applied representation is discussed in terms of impact on automated segmentation algorithms and known advantages of medial models are revisited.

Keywords: inner ear, segmentation, active shape models, cochlea

1. INTRODUCTION

The concept of Active Shape Models (ASM) is widely used for medical image analyzing.¹ A major advantage of statistical shape based techniques, like ASM, is the possibility to identify relevant structures in target image data with certain expert knowledge. This knowledge is generated by training data for the deformable model. It is possible to enhance target data by additional information which is not part of the image data, e.g. positions of important membranes of the cochlea and connected structures. The visibility of these membranes strongly depends on the image data modality and might be only identifiable by experts. This technique even applies to distinct structures like the cochlea with its helical shape. An application is a planning system for insertion of cochlear implants (CI).² Noble et al. suggest a method to represent the cochlea based on Point Distribution Model (PDM)³ and recently a similar approach using Probabilistic Active Shape Models (PASM) was published by Becker et al.⁴

Our work describes the application of a medial (skeleton) representation. The cochlear tubular structure is further examined, and a generic way to build a medial model is constituted. The cochlear structure suits well to profit from multiple advantages of medial models.⁵ One realization of medial representation is called *m-reps* and was introduced by Pizer et al.⁶ Several versions, enhancements and applications of *m-reps* have been published so far.^{7,18} Segmentation of tubular structures was introduced by Saboo et al.⁸ In later work it was pointed out that the need for an appropriate initial design for a *m-rep* is one of the major disadvantages of this medial representation due to the difficult ratio between quality and grade of compactness.⁷ While in many cases the *m-rep* form in 3D is based on a grid or a lattice of so called *atoms*, it will be demonstrated that the serial linking of the atoms is an adequate base for the tubular structure of the cochlea.

Since the underlying construct of the model is simply a chain with no branching, the aspect of correspondence between different models is handled easily. *Hubs* describe the position of an atom on the center line. The model defines its own coordinate system which allows finding the correspondence between all used *hubs* by examining their distances to predefined *fiducial hubs*. Such a *fiducial hub* can be the position on the centerline of the cochlea closest to the round or oval window. This also leads to a more far-reaching advantage, e.g. choosing a length

Further author information: Johannes Gaa

E-mail: johannes.gaa@imes.uni-hannover.de, Telephone: +49 (0)511 762 19560

of cochlear implant electrode or customize cochlear implant strategies.⁹ Another advantage of medial models is the straight forwardness for interaction. An appropriately designed medial model allows modifying its bounding by verifying each atom independently. This characteristic allows interaction to correct or enhance the results of automated segmentation algorithms.

2. METHODS

The medial representation of the cochlea is discrete. Thus, several atoms along the skeleton line of the cochlea describe the shape. Each atom is defined by $\mathbf{m} = \{\mathbf{x}, r, \vec{n}_0, \vec{n}_1\}$. Thereby, an atom consists of the *hub* $\mathbf{x} \in \mathbb{R}^3$, the length $r \in \mathbb{R}$ of the *spokes* and the two directions of the spoke $\vec{n}_0, \vec{n}_1 \in \mathbf{S}^2$ defined on the unit sphere. Figure 1 illustrates the structure in training data. Additionally, the atom elements, its frame (it is possible to define m-reps using the frame instead of the spoke directions) and an orientation in each atom is shown. The orientation of the atom is used to span the search space for the spokes correctly. This step is further described in 2.1.

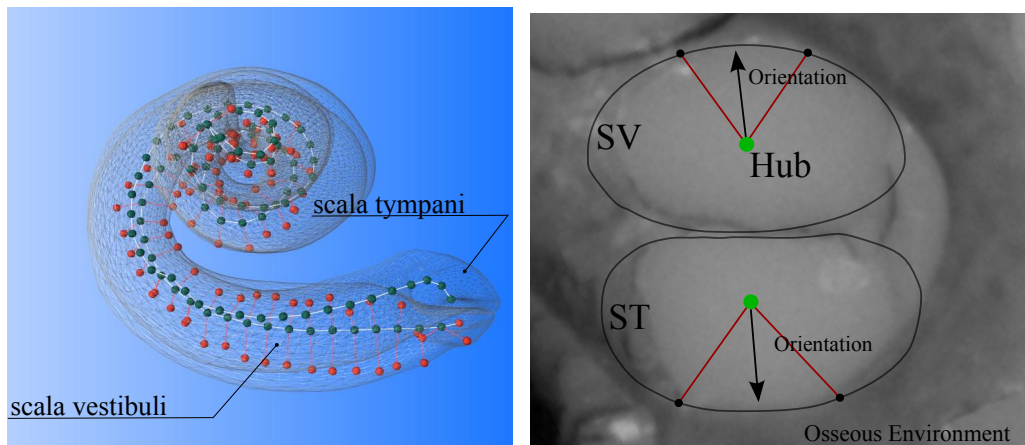


Figure 1. (left) *m-rep* extracted from a trainings shape, assembled by hubs (green) and their bounding describing spokes (red). Only the spokes in scala vestibuli are shown for better of overview. (right) The atom orientation point towards the osseous environment.

As direct product of atom elements, m-reps are more complex and common Euclidean actions are invalid. In particular the principal component analysis (PCA) as core function of statistical shape models can not be applied. However, Fletcher et al. described how to use the more general pendant to the PCA, called principal geodesic analysis (PGA) for generating deformable models based on geodesics.¹⁰ We implemented PGA to generate a deformable model for the cochlea.

Since our training data consists of manually segmented shapes, their representation is landmark based. This gives the opportunity to align the training data using the *Generalized Procrustes* algorithm¹¹ and to avoid the more complex alignment in medial form. On the down side a method to convert the point based shapes to medial representation is needed. A roughly outlined work-flow to generate a deformable medial model contains:

- Aligning the training data using Generalized Procrustes algorithm,
- converting each training shape in medial form by extracting the medial loci and creating atomic structure,
- and generating a deformable model with the variance within the converted training data.

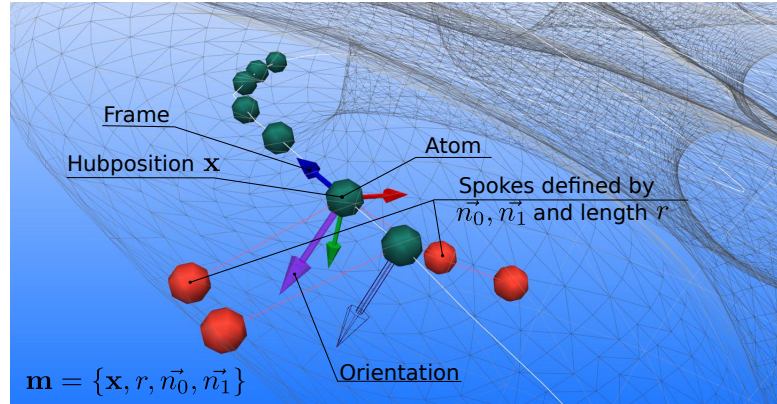


Figure 2. All components of a single (internal) atom to describe a region: Hub position on center line, two spokes to a bounding point, the frame in the atom and the orientation.

2.1 Preparing the Training Data

The results of a model based segmentation rely strongly on the used training data. For this work four manually segmented shapes served as trainings data, each representing *scala vestibuli* and *scala tympani* as well as the basilar membrane between. The shapes were segmented on micro-grinding image data¹² to extract the information of the membrane. A pre-generated template mesh was manually fitted into each training image data to preserve the point correspondence between each training shape.

The first step is to convert the training shape to a accurate m-rep form. Thus, a skeleton of each input shape needs to be calculated. Different techniques to achieve this were discussed in earlier work. Siddiqi et al. describe selected methods of thinning- and skeletization-algorithms in the context of medial models.⁵ While they also consider two-dimensional methods, the application in three-dimensional images is crucial for this work.¹³ Our function is a direct derivative of the method suggested by Antiga et al.¹⁴

Once the center line of the data is defined, it is examined for three fiducial points. The cochlea is divided into the two *scalae* by finding the closest hubs to the oval window, the round window and the apex. This can be automated by finding the highest curvature κ_{max} along the center line respectively to the length of the arc. One can examine, for instance the tangents in discrete points on the center line. This point represents the passage from *scala tympani* to *scala vestibuli*. However, depending on the method to find the center line, we interpolate it with a spline, to ensure the detection of the highest curvature at the apex. If necessary, each hub is identify by the distances either to the apex hub and to the oval window hub or to the apex and the round window hub. After identification the correspondence between the atoms in the population is setup.

As mentioned before one particular aspect when specifying the spokes for the atoms is the orientation of the atom. The thin membranes within the cochlea have a very weak appearance in common medical image data (like CT), if there is any at all. With this in mind, we set the orientation during the identification of the spokes towards the usually strong appearing bounding in the osseous environment. The mean orientation of all casting rays $\vec{r}_i \in \mathbb{R}^3$, $\vec{r}_i \perp \vec{b}_t$ where \vec{b}_t is the tangential base vector in each atom \mathbf{m} is determined. Although, only the casting rays which do not intersect drastically with the membranes are considered. By rotating the summation of all those vectors \vec{t} by $\pm 60^\circ$ the search subspace for possible spokes is defined. Finally we link the complete atom structures to each other to build up the m-rep.

2.2 Deformation of the medial model

The medial description and deformation of a model using PGA is based on the mapping between the submanifold and the tangential space.¹⁰ We follow the notation introduced by Fletcher et al. for the exp map $Exp_p(u)$ and the log map $Log_p(m)$. Depending on the input, PGA allows us to find the deformation parameters v_k , very similar as PCA for ASM. With the intrinsic mean μ of all N inputs we use the method to find the principal directions as eigenvectors v_k of

$$\vec{S} = \left(\frac{1}{N} \right) \sum_{i=1}^N (u_i u_i^T), \quad (1)$$

where u_i describes the mapped shape x_i , $\text{Log}_\mu(x_i)$.

A new representation can be generated with in the statistical restrictions of the training data by varying the parameters b_k in

$$M(b) = \exp_\mu \left(\sum_{k=1}^d (b_k v_k) \right). \quad (2)$$

Preparing the four input shapes and extracting mean shape as well as the shape parameter required an average runtime of 4.7 seconds on a workstation (Kernel Linux 3.2.0-23-generic, CPU: Intel® Core™ i5-2500 CPU @ 3.30 GHz).

The actual fitting of the presented medial model of the cochlea into a specific image data can be more complex due to the high number of influencing factors. For instance, the initial positioning of the model in the data set has usually a high impact on the result and computation time. Wimmer et al. recently published a method to quickly place an ASM in the image by manually defining the cochlear axis.¹⁵ In the case of the medial models this issue can be traced down to a spline fitting problem, representing the medial locus of the model. Several methods regarding this have been discussed, i.e. by Akemi and Andrés¹⁶ and more general by de Bruijne et al.¹⁷ This work focuses only on the higher scaled fitting of the model parameters b_k introduced in Equation 2. We leave a further examination of the pose initialization based on a medial representation for a later publication.

Assuming the model is good enough placed, the optimization of the parameters b_k is done. The cost function $f_j(I, u)$ is minimizing the difference of the average appearance in the target image data at each spoke end and the mean appearance based on the trainings image data. This appearance is predefined in the a region of interest (ROI) of the image. is

$$\vec{x}_j = \arg \min_{\vec{u}} f_j(I, \vec{u}), \quad (3)$$

where \vec{u} represents the position of the spokes end. The result of this optimization is the vector b_k , which deforms the model so the spokes fit best to the boundary found in the image. Algorithm 1 describes the basic steps.

Algorithm 1: Optimization

Input: medial model deformation vector b_k , target image data
while (result not found):
 vary elements of vector b_k ;
 for each atom of model:
 compare ROI of target image at spoke end with mean ROI;
 evaluate fitness;

3. RESULTS

This section describes the application of a generated medial model for image segmentation. The model was extracted from four manually segmented trainings data as mentioned in section 2.1, each training shape sectioning *scala tympani*, *scala vestibuli* and apex. The cochlear medial models were automatically generated, identifying hub position, spoke directions, and the orientation towards the osseous environment in each atom. The number of atoms was pre-set at 100. After identifying the atoms describing the passage from one *scala* to the other in each training shape, the atom correspondence were automatically found using the distances to apex, start and end of the atom chain. Based on that the variance through the population was determined using PGA and the deformable model was generated. The initial pose of the mean shape was identified by a point based registration method, using the hub positions of apex-, start- and end-atom and the same landmarks manually picked in the image data.

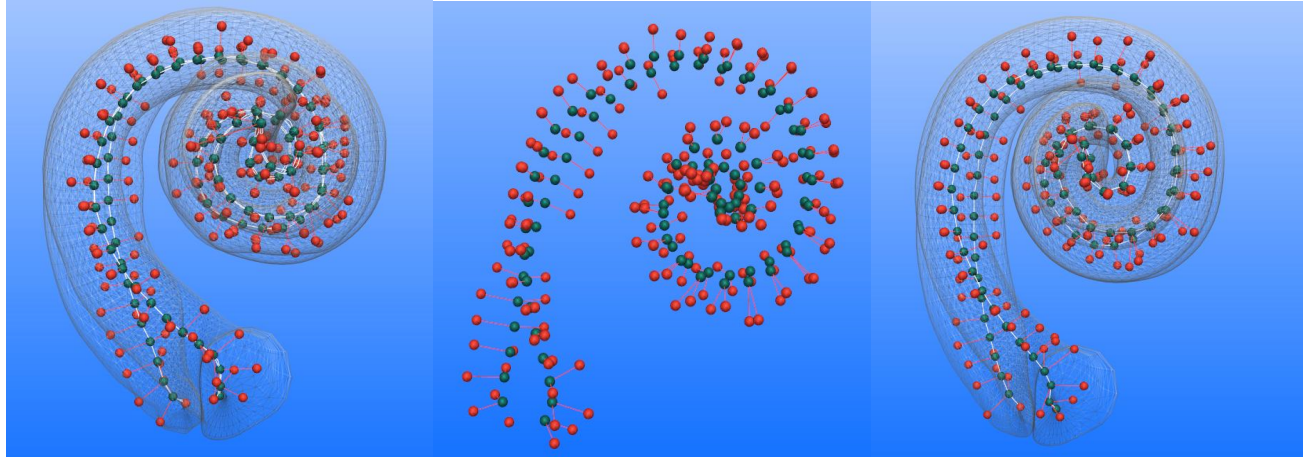


Figure 3. The image in the center displays the mean representation used for the initialization and segmentation in the patient image data. To the left and to the right two resulting shapes M_{σ_1} and M_{σ_2} with different parameter sets b_{σ_1} and b_{σ_2} are shown. They represent two of the trainings shapes used for the deformational model.

Figure 3 illustrates the mean model of with two of its trainings shapes. Due, to the low resolution of image data the manually created segmentation often represents the complete cochlea as tubular structure, while our medial model approach gives an approximated position for each scala of the cochlea. This approximation is based on the position specified during the segmentation of the trainings data.

An exemplary segmentation of a cone beam computer tomography (CBCT) image data set was done to get a first impression of run time and feasibility (see Tab. 1).

4. DISCUSSION AND CONCLUSION

In this study, the usability of medial models for cochlea segmentation was evaluated. A deformable model inspired by the common m-reps was generated. Therefore, a small number of training data was prepared and converted to the necessary medial representation. It was shown that this conversion can be done automatically. While several methods for (automated) cochlear analyzing have been published, to the best of our knowledge, this report presents the first usage of medial models. Modeling the cochlea as a quasi-tubular structure implies new methods to determine patient specific inner ear characteristics, i.e. the length of *scala tympani*. Due to the small number of trainings data the model leaks variance in areas. This can be resolved using a higher number of trainings data. The deformable model was finally used to segment an exemplary computed tomography image. As future work a more extensive analysis of segmentation results using medial models is planned.

Medial models haven been successfully applied for segmenting tubular structures, i.e. by Pouch et al.¹⁸ Due to its reduced solution space, the model promises short optimization time, which also has to be evaluated in future work.

Process Step	CBCT
	Data Set 1
Initialization [ms]	655
Segmentation [ms]	7344

Table 1. Calculation and runtime for an exemplary CBCT image data.

ACKNOWLEDGMENTS

This work was funded by the German Federal Ministry of Education and Research (BMBF). The project number is 13GW0019C. Responsibility for the contents of this publication lies with the authors.

REFERENCES

- [1] Cootes, T. F., Taylor, C. J., Cooper, D. H., and Graham, J., "Active shape models and their training and application," *Comput. Vis. Image Underst.* **61**(1), 38–59 (1995).
- [2] Noble, J. H., Dawant, B. M., Warren, F. M., and Labadie, R. F., "Automatic identification and 3D rendering of temporal bone anatomy," *Otol. Neurotol.* **30**(4), 436–442 (2009).
- [3] Noble, J. H., Gifford, R. H., Labadie, R. F., and Dawant, B. M., "Statistical shape model segmentation and frequency mapping of cochlear implant stimulation targets in CT," *Med Image Comput Comput Assist Interv* **15**(Pt 2), 421–428 (2012).
- [4] Becker, M., Kirschner, M., and Sakas, G., "Segmentation of risk structures for otologic surgery using the probabilistic active shape model (pasm)," *Proc. SPIE* **9036**, 90360O–90360O–7 (2014).
- [5] Siddiqi, K. and Pizer, S., [*Medial Representations: Mathematics, Algorithms and Applications*], Springer Publishing Company, 1st ed. (2008).
- [6] Pizer, S., Fletcher, P., Joshi, S., Thall, A., Chen, J., Fridman, Y., Fritsch, D., Gash, A., Glotzer, J., Jiroutek, M., Lu, C., Muller, K., Tracton, G., Yushkevich, P., and Chaney, E., "Deformable m-reps for 3D medical image segmentation," *International Journal of Computer Vision* **55**(2-3), 85–106 (2003).
- [7] Yushkevich, P., Zhang, H., and Gee, J., "Continuous medial representation for anatomical structures," *Medical Imaging, IEEE Transactions on* **25**(12), 1547–1564 (2006).
- [8] Saboo, R. R., Rosenman, J. G., and Villarruel, C., "Segmentation of tubular objects by posterior optimization of m-reps," *UNC Chapel Hill technical report* (2006).
- [9] Noble, J., Labadie, R., Gifford, R., and Dawant, B., "Image-guidance enables new methods for customizing cochlear implant stimulation strategies," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on* **21**(5), 820–829 (2013).
- [10] Fletcher, P. T., Lu, C., Pizer, S. M., and Joshi, S. C., "Principal geodesic analysis for the study of nonlinear statistics of shape," *IEEE Trans. Med. Imaging* **23**(8), 995–1005 (2004).
- [11] Goodall, C., "Procrustes methods in the statistical analysis of shape," *Journal of the Royal Statistical Society. Series B (Methodological)* **53**(2), 285–339 (1991).
- [12] Rau, T., Würfel, W., Lenarz, T., and Majdani, O., "Three-dimensional histological specimen preparation for accurate imaging and spatial reconstruction of the middle and inner ear," *International Journal of Computer Assisted Radiology and Surgery* **8**(4), 481–509 (2013).
- [13] Leymarie, F. F., Kimia, B. B., and Giblin, P. J., "Towards surface regularization via medial axis transitions," in [*Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*], 123–126 (2004).
- [14] Antiga, L., Iordache, E. B., and Remuzzi, A., "Computational geometry for patient-specific reconstruction and meshing of blood vessels from MR and CT angiography," *IEEE Transactions on Medical Imaging* **22**(5), 674–684 (2003).
- [15] Wimmer, W., Venail, F., Williamson, T., Akkari, M., Gerber, N., Weber, S., Caversaccio, M., Uziel, A., and Bell, B., "Semiautomatic cochleostomy target and insertion trajectory planning for minimally invasive cochlear implantation," *BioMed Research International* , 8 (2014).
- [16] Akemi, G. and Andrés, I., "Firefly algorithm for explicit b-spline curve fitting to data points," *Mathematical Problems in Engineering* **2013**(4), 12 (2013).
- [17] de Bruijne, M., van Ginneken, B., Viergever, M., and Niessen, W., "Adapting active shape models for 3D segmentation of tubular structures in medical images," *Information Processing in Medical Imaging* **2732**, 136–147 (2003).
- [18] Pouch, A. M., Yushkevich, P. A., Jackson, B. M., Gorman III, J. H., Gorman, R. C., and Sehgal, C. M., "Dynamic shape modeling of the mitral valve from real-time 3d ultrasound images using continuous medial representation," *Proc. SPIE* **8320**, 832009–832009–8 (2012).